WL-TR-93-1011

MULTIRESOLUTION ANALYSIS OF SAR DATA

ROBERT HUMMEL

NEW YORK UNIVERSITY
COURANT INSTITUTE OF MATH. SCIENCES
251 MERCER STREET
NEW YORK NY 10012

JAN 1993

FINAL REPORT FOR 09/01/89-09/30/92

AVIONICS DIRECTORATE
WRIGHT LABORATORY
AIR FORCE MATERIEL COMMAND
WRIGHT PATTERSON AFB OH 45433-7001

NOTICE

WHEN GOVERNMENT DRAWINGS, SPECIFICATIONS, OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY GOVERNMENT-RELATED PROCUREMENT, THE UNITED STATES GOVERNMENT INCURS NO RESPONSIBILITY OR ANY OBLIGATION WHATSOEVER. THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA, IS NOT TO BE REGARDED BY IMPLICATION, OR OTHERWISE IN any MANNER CONSTRUED, AS LICENSING THE HOLDER, OR ANY OTHER PERSON OR CORPORATION; OR AS CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE, OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

THIS REPORT IS RELEASABLE TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS). AT NTIS, IT WILL BE AVAILABLE TO THE GENERAL PUBLIC, INCLUDING FOREIGN NATIONS.

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

Martin E. Justice

MARTIN E. JUSTICE
Project Engineer

James C. Rachal

JAMES C. RACHAL, CHIEF
Development Section
Target Recognition Technology Branch

FOR THE COMMANDER

Edmund G. Zelnio

EDMUND G. ZELNIO, Actg Chief
Mission Avionics Division
Avionics Directorate

IF YOUR ADDRESS HAS CHANGED, IF YOU WISH TO BE REMOVED FROM OUR MAILING LIST, OR IF THE ADDRESSEE IS NO LONGER EMPLOYED BY YOUR ORGANIZATION, PLEASE NOTIFY WL/AARA BLDG 23, 2010 FIFTH ST., WRIGHT-PATTERSON AFB, OH 45433-7001 TO HELP MAINTAIN A CURRENT MAILING LIST.

COPIES OF THIS REPORT SHOULD NOT BE RETURNED UNLESS RETURN IS REQUIRED BY SECURITY CONSIDERATIONS, CONTRACTUAL OBLIGATIONS, OR NOTICE ON A SPECIFIC DOCUMENT.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	JAN 93	Final, Sep 89-Sep 92

4. TITLE AND SUBTITLE
Multiresolution Analysis of SAR Data

5. FUNDING NUMBERS
C: F33615-89-C-1087
PE: 62204F
PR: 7629
TA: 04
WU: 08

6. AUTHOR(S)
Robert Hummel

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
New York University
Courant Institute of Mathematical Sciences
251 Mercer Street
New York NY 10012

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AVIONICS DIRECTORATE
WRIGHT LABORATORY (WL/AARA, 2010 Fifth St)
AIR FORCE MATERIEL COMMAND
WPAFB OH 45433-7001
ATTN: MARTIN E. JUSTICE (513-255-1109)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
WL-TR-93-1011

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is limited.

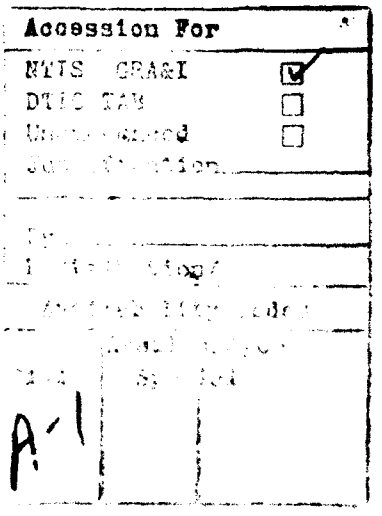
12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The "Multiresolution Analysis of SAR Data" program has supported research work in five areas: (1) Geometric hashing theory can now be viewed as a Bayesian approach to object recognition. False alarm rates can be greatly reduced by using certain enhancements and modifications developed under this project. (2) Geometric hashing algorithms now exist for the Connection Machine. Recognition of synthetically-produced dot arrays has been demonstrated using a model base of 1024 objects. The work represents a substantial advance over existing model-based vision capabilities. (3) Algorithms have been developed for determining the translation and rotation of a sensor given only the image flow field data. These are new algorithms, and are much more stable than existing computer vision algorithms for this task. The algorithms might provide independent verification of gyroscopic data, or might be used to compute relative motion with respect to a moving scene object, or may be useful for motion-based segmentation. (4) Our theories explaining the Dempster/Shafer calculus and developing new uncertainty reasoning calculi have been extended, and presented at a conference and were incorporated into the Bayesian interpretation of geometric hashing (see (1) above). (5) "Wavelet Slice Theorem" has been developed in several different versions, any of which yields an alternate approach to image formation. The result may well provide a more stable approach to image formation than the standard Fourier-based projection slide theorem, since interpolation of unknown spectral values is better-founded.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Geometric hashing, uncertainty reasoning, wavelets	20
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Table of Contents



List of Figures

Multiresolution analysis of SAR Data

Final Report

Robert Hummel
New York University

1. Overview

The "Multiresolution Analysis of SAR Data" project (Contract F33615-89-1087, reference (BAA) 87-02-PMRE), has supported research work in five areas:

(1) Geometric hashing foundations and theory for model-based vision;

(2) Parallel implementations of geometric hashing;

(3) Motion parameter estimation algorithms for egomotion;

(4) Uncertainty reasoning and evidential reasoning studies; and

(5) The use of wavelets for SAR image construction.

All of the items except (3) are reasonably part of the original SOW. Item (3) developed due to the work of V. Sundareswaran, who was supported by this grant after I. Rigoutsos won support from an IBM Fellowship.

The following are the principal advances that have been made in each of these areas:

(1) Geometric hashing theory can now be viewed as a Bayesian approach to object recognition. False alarm rates can be greatly reduced by using certain enhancements and modifications developed under this project.

(2) Geometric hashing algorithms now exist for the Connection Machine. Recognition of synthetically-produced dot arrays has been demonstrated using a model base of 1024 objects. The work represents a substantial advance over existing model-based vision capabilities.

(3) We have developed new algorithms for determining the translation and rotation of a sensor given only the image flow field data. These are new algorithms, and are much more stable than existing computer vision algorithms for this task. The work has been accepted for publication in a major journal. The algorithms might provide independent verification of gyroscopic data, or might be used to compute relative motion with respect to a moving scene object, or may be useful for motion-based segmentation.

(4) Our theories explaining the Dempster/Shafer calculus and developing new uncertainty reasoning calculi have been extended, and presented at a conference, and were incorporated into the Bayesian interpretation of geometric hashing (see (1) above).

(5) We have developed a "Wavelet Slice Theorem," in several different versions, any of which yields an alternate approach to image formation. The result may well provide a more stable approach to image formation than the standard Fourier-based projection slice theorem, since interpolation of unknown spectral values is better-founded. Moreover, we have made considerable progress on an alternative approach to SAR image formation that avoids the computationally expensive steps of rerasterizing return data, and Fourier transforms, using instead a more complex chirp function. The method suggests a simple real-time implementation that builds up the final image incrementally, during the fly-by.

The list of significant technical reports and publications produced as a result of the research support of this grant are listed below.

[1] R. Hummel, "Geometric Hashing Methods for Object Recognition," NYU Proposal to DARPA in response to BAA 90-15, submitted October 1990, reviewed and approved (for future funding),

January, 1991.

[2] I. Rigoutsos and R. Hummel, "Several Results on Affine Geometric Hashing," *Proceedings of the 8th Israeli Conference on Artificial Intelligence and Computer Vision*, pp. 1-12, December 1991, Tel Aviv, Israel.

[3] I. Rigoutsos and R. Hummel, "A Bayesian Approach to Model Matching with Geometric Hashing," NYU Technical Report, 1991, available by anonymous ftp from cs.nyu.edu under /pub/local/hummel/papers/bayeshashing.

[4] I. Rigoutsos and R. Hummel, "Robust Similarity Invariant Matching in the Presence of Noise: A Data Parallel Approach," *Proceedings 8th Israeli Conference on Artificial Intelligence and Computer Vision*, pp. 27-43, December 1991, Tel Aviv, Israel.

[5] Y. Lamdan and H. Wolfson, "On the Error Analysis of Geometric Hashing," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **CVPR91**, 1991, pp. 22-27.

[6] H. Wolfson and Y. Lamdan, "Transformation Invariant Indexing," in *Invariance in Computer Vision*, (Eds. J. Mundy and A. Zisserman), pp. 335-353, MIT Press, 1992.

[7] H. Wolfson, "Articulated Object Recognition, or, How to Generalize the Generalized Hough Transform," (with A. Beinglass), *Proceeedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 461-466, June 1991.

[8] I. Rigoutsos and R. Hummel "Massively Parallel Model Matching: Geometric Hashing on the Connection Machine," *IEEE Computer*, pp. 33-42, February 1992.

[9] I. Rigoutsos and R. Hummel, "Implementation of Geometric Hashing on the Connection Machine," *Proceedings IEEE Workshop on Directions in Automated CAD-based Vision Systems*, June 1991, Maui, Hawaii.

[10] I. Rigoutsos and R. Hummel, "On a Parallel Implementation of Geometric Hashing on the Connection Machine," CIMS Technical Report #554, April 1991 (Extended version of the previous paper).

[11] R. Hummel and V. Sundareswaran, "Motion estimation using global image flow velocities," ARVO Abstract, 1991.

[12] V. Sundareswaran, "Egomotion from global flow field data," *Proceedings of the Workshop on Image Motion Analysis*, Princeton, N.J., October, 1991.

[13] R. Hummel and V. Sundareswaran, "Motion parameters from global flow field data," To Appear in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, May 1993.

[14] V. Sundareswaran, "A fast method to estimate sensor translation," in *Proceedings of the Second European Conference on Computer Vision*, Santa Margherita, Italy, May 1992.

[15] R. Hummel and L. Manevitz, "Statistical approaches to fusion with uncertainty," Session MB4, *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Charlottsville, Virginia, October, 1991.

We next give technical details on the progress in each area.

2. Geometric Hashing

Geometric hashing is an approach to model-based object recognition that is analogous to the hashing approach to data retrieval in computer science. The idea is that collections of object features are encoded in such a way that recognition of a few features may enable indexing into a table that then designates a list of possible objects. The intent is to reduce the complexity of search processing by allowing scene features to index into the model database, rather than positing matches between scene features and model features and then searching for a hypothesis that can be verified.

The initial idea is due to Jack Schwartz, who applied the idea to curve matching for robotic applications. In this form, the idea was related to, but somewhat dissimilar from, the idea of R-tables in the

generalized Hough transform. Subsequently, Haim Wolfson and his students, working with Professor Schwartz, extended the idea to other object-recognition tasks, including 3-D curve matching, affine-invariant object matching based on point sets, affine-invariant matching based on curves using concavities as primitive features, matching using convex curve portions for recognition, and matching of objects based on line extractions. A milestone was reached with the publication of the thesis of Y. Lamdan, at NYU, who studied hashing methods for 3-D object recognition from 2-D projections. His thesis is recommended reading for an understanding of the basis of the idea and the many variations and possibilities. The idea has been taken up by other groups, and has been studied by researchers at the University of Washington, MIT, Syracuse University, and in Europe, especially at INRIA-Rocqencourt.

With that background, and support from AFAL, we have subsequently developed the theory of geometric hashing, with the following intents:

(1) We have analyzed the stability and false alarm rates, both experimentally and theoretically;

(2) Concentrating on rigid-transformation and similarity-transformation invariance, we have implemented systems using thousands of models of dot patterns;

(3) In order to improve the efficiency of the method, we have developed a number of enhancements, including (1) Hash table equalization, (2) Improved hash functions and use of symmetries in the hash table, and (3) folding of the hash table.

(4) We have developed a Bayesian interpretation of geometric hashing that permits a well-founded basis for weighted voting in the evidence-gathering phase of geometric hashing.

(5) We have developed the application of geometric hashing to articulated object recognition, which should facilitate recognition of models which can vary according to one or a few parameters.

All of this work is described in detail in the publications [1-7] (from the list of references in Section 1).

The application of geometric hashing to ATR problems, both with SAR data and IR sensor data, is clear. However, we have not pursued these issues extensively in this project, because the development of such applications is critically dependent on stable feature extraction methods. Such algorithms are being extensively developed by DoD contractors, but require large databases for evaluation. Much of our work has used synthetic data, based on point feature sets, with noise added, in order to evaluate the power of the indexing methods. Recently, we have begun to experiment with recognition of military aircraft from silhouettes and extracted edge features. However, even in this application, our data comes from freely-available pictures and sketches. Rigoutsos will also use images of automobiles in his continuing thesis work. These objects are somewhat more difficult because of their more rotationally symmetric shapes. That is, the military aircraft tend to be elongated with distinguishing marks. To date, we achieve extremely good results for similarity recognition of military aircraft using side-views, employing a standard edge detector and automatic point extraction.

Most of the developments are incremental, with the possible exception of the Bayesian-interpretation work, which might be considered more revolutionary. Accordingly, we explain the Bayesian interpretation in general terms here.

In similarity-invariant matching of point sets using geometric hashing (which we use here as an example), pairs of points are chosen from the scene as a basis pair. For each basis pair, a coordinate system is defined, and the other scene points can be expressed as coordinate pairs in any given coordinate system. For a given basis, each point thereby hashes to a location in a two-dimensional hash space. In classical geometric hashing, as it has been described heretofore, the hash location is regarded as landing within a bin, and each entry that landed within that bin during a preprocessing phase receives a vote. However, during the preprocessing phase, each entry is associated with a position within the two-dimensional hash space. Rather than regarding the entries as belonging to bins, we regard them as fixed points in the hash space. Accordingly, when a point hashes into the space, it should vote for *nearby* entries. Further, it seems reasonable that the weighted degree of vote should depend on the distance

between the hash point and the entry. The larger the distance, the smaller the vote should be. The rate of fall-off of the vote should depend on the expected degree of noise in the location of the hash points, which in turn depends on the expected noise in the placement of the points in the scene, and in the distance between the points in the basis pair. When one applies a strict Bayesian analysis to the situation, the result is a precise formula for the weighted contribution that a hash to location (u,v) should make to an entry for model k using basis (i,j) located at a point (x,y) in hash space. The contribution depends on the standard deviation of the noise σ in the scene, and the basis separation D. In this case, the contribution is:

$$z = \frac{(u^2+v^2+3)^2}{(4(x^2+y^2)+3)\cdot\sigma^2/D^2}\cdot\exp\left[\frac{-\|(u,v)-(x,y)\|^2}{(4(x^2+y^2)+3)\cdot\sigma^2/D^2}\right].$$

Figure 1 shows an example of recognition of a military aircraft from a database of 15 aircraft. The scene comes from a grayscale image of one of the aircraft types, but is not taken from the same data that are used to construct the model database. **The points in the scene are extracted automatically in this example.** The results are easily extended to affine invariant matching.

We believe that this work forms the leading-edge of the object recognition research domain, as was evidenced by the results presented at the 1991 Workshop on Automated CAD-based vision systems. The theoretical advantages of the geometric hashing method are clear, and the issue remains as to verify the applicability through experiments and through refinements. Most other researchers are working with databases consisting of one or two objects, and most work remains in the realm of geometric reasoning and hypothesis-and-verify. Some groups have begun using geometric hashing methods and their variants, and we have benefited from their work, but the NYU work remains more advanced in terms of both applications and generality.

3. Parallel algorithms for geometric hashing

A major advantage of the geometric hashing approach to object recognition is its parallelizability. However, there are many different ways to parallelize the algorithm. Chosing an appropriate parallel implementation can be very subtle.

Rigoutsos, as part of his thesis work, has implemented a number of versions of geometric hashing for execution on the Thinking Machines Inc. Connection Machine™. We have gained access to various Connection Machines by making use of the DARPA "Network Server" program. He is currently completing an updated implementation of one of the versions which will support the weighted voting as described in the Bayesian interpretation above, and will be used for the rapid recognition of military aircraft and automobiles.

As is very typical, there is an interesting and subtle parallel approach which has good asymptotics, and there is a much simpler and less sophisticated algorithm that is likely to present the more practical approach in general. The sophisticated approach uses a "connectionist" view of geometric hashing, whereas the simpler version is based on a "hash location broadcast" model. The two approaches are best understood from the paper that will appear in *IEEE Computer* (Reference [8]). This paper, which will be a very high-profile presentation of the geometric hashing paradigm and the parallelizability of the methods, is a relatively short and "popular" description, intended for a very broad audience (all members of the IEEE Computer Society receive the journal). References [9] and [10] further describe aspects of the parallel implementation.

The fundamental choice in our parallelizations of geometric hashing is to continue to serialize over choices of basis pairs in the scene, but to maintain the entire hash table data base in the memory of the parallel machine, and to make sure that the search over model/basis pairs is done in parallel. This choice is motivated by the fact that the advantage of geometric hashing is that it makes the search over model/basis pairs fast, whereas it retains the need to search over the image for a basis set. However logical this choice seems, all other parallel implementations of geometric hashing of which we are aware

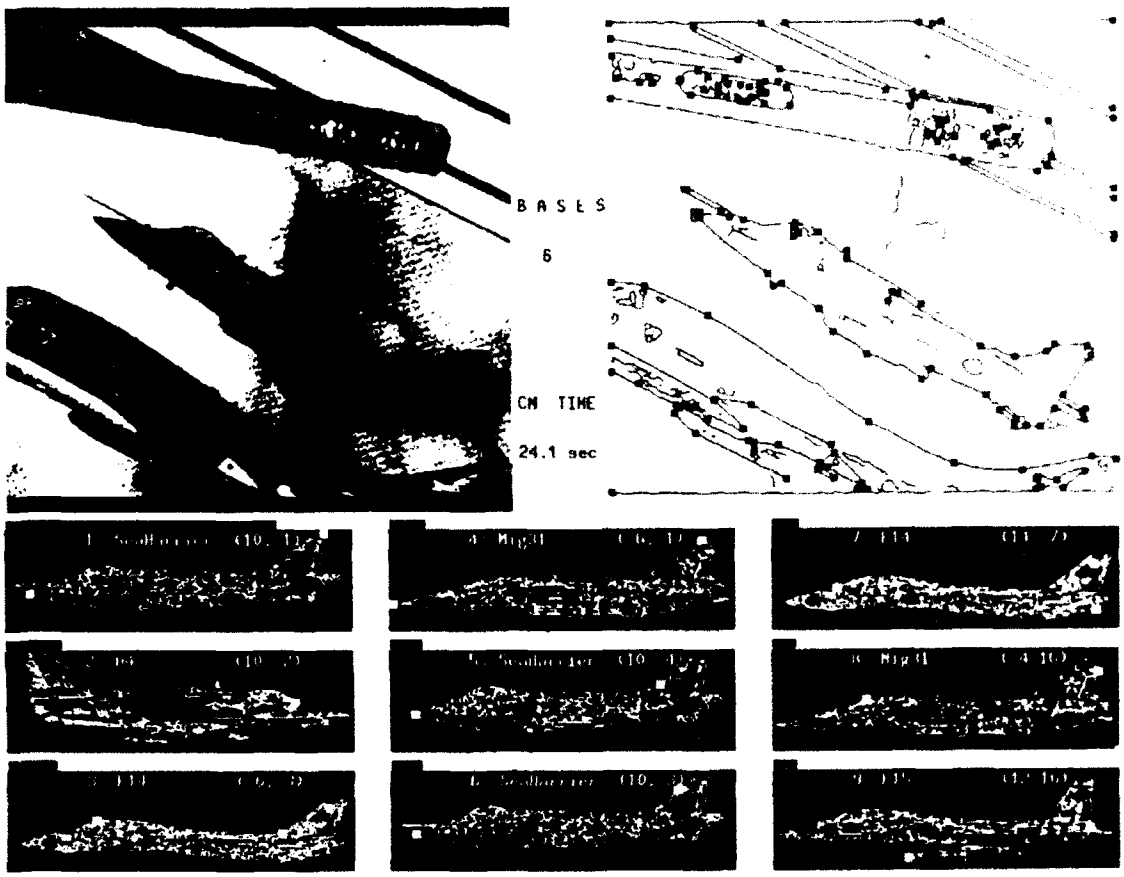


Figure 1. Recognition of a military aircraft. The database consists of 15 military aircraft types, where points were selected manually from scanned images of side-view drawings of each aircraft. The scene is a photograph of one of the aircraft types, in flight. The photograph comes from a different book than the sketches that are used for the model base. Edges and points are extracted automatically from the scene, and the Bayesian version of the geometric hashing method is used for recognition. The nine top model/basis pairs are shown; the one receiving the greatest weight is the correct recognition of the aircraft. Bars over the sketches indicate the degree of support for each model/basis.

(including the Martin Marietta Denver approach as of last summer), parallelize over a different data set.

Allow us to explain in somewhat greater detail the broadcast approach. We have, from the preprocessing phase, a large number of records of the form (M_k,B,l,x,y,C), where M_k is a model number, B is a basis within the model, l is another point in the model (not in the basis), (x,y) is the location to which the point hashes under the basis B, and C is a covariance matrix which, based on expected noise of placement of points in the scene, predicts the covariance of the distribution of the location of the hash value of the point l about the location (x,y) under the basis B. The hash table is no longer organized as a table, but rather as simply a large collection of such records.

During the recognition phase, a basis is chosen and fixed. We operate sequentially over the other points in the scene. It is this degree of serialism that one would like to parallelize, and leads to the more complicated connectionism algorithm. In the broadcast algorithm, each point in the scene successively is hashed to a location (u,v) in the hash table, and the location (u,v) is broadcast to all the records stored in the hash table. (Some savings can be obtained by broadcasting to only those records that lie near (u,v), which means that the records must be organized in a suitable manner in advance.) Each record computes the distance between its (x,y) and the location (u,v), and uses the predicted covariance C to compute the degree of contribution that the scene point should lend to the associated information (M_k,B,I). Votes are accumulated and after cycling through all appropriate scene points, votes are summed for model/basis pairs (M_k,B). These form the desired evidence values.

Figure 2 shows an example of a recognition that takes place using the broadcast algorithm, as implemented on a 16K-processor Connection Machine.

Many variations are possible. However, as stated, the algorithm is able to work efficiently and rapidly on an 8K-processor CM-2 with 16 models, with approximately 20 points in each model, using scenes of roughly 100 points. Currently, each probe takes about 2 seconds, and typically 30 probes are needed (i.e., 30 choices of basis pairs) in order to find the model. Approximately linear speedup can be achieved on a larger Connection Machine, (e.g., a 64K model), either a much more complicated algorithm must be used, or the problem must be scaled up. The current implementation on the 8K-processor machine can handle 32 models in the database without any degradation in processing time.

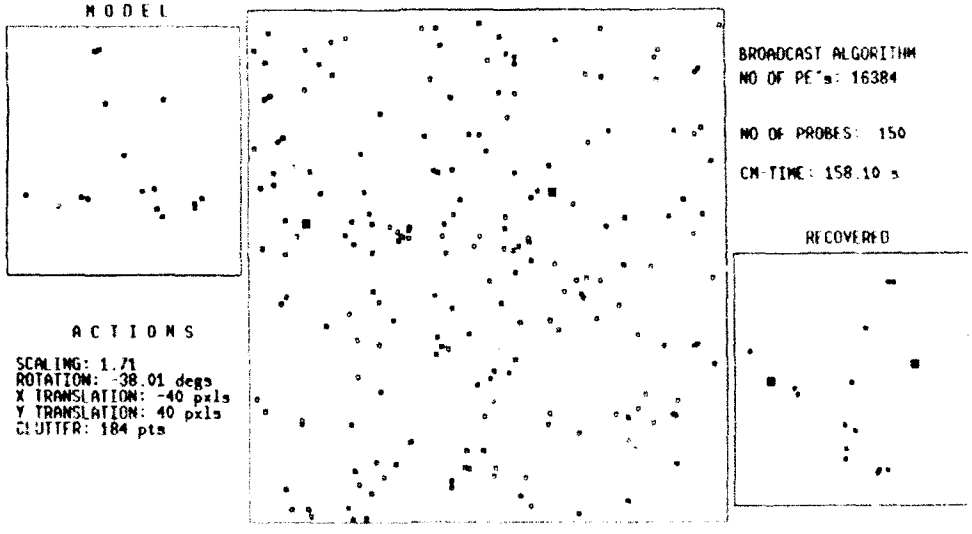


Figure 2. An example recognition using the broadcast algorithm on a Connection Machine. There are 1024 models in the database. Each model consists of 16 points. The scene contains 200 points. Approximate 0.8 second was required for each probe on a 16K-processor CM. By using symmetries in the hash table, this time could easily be halved. A roughly fourfold speedup would be expected by using a 64K-processor CM. For this example, 150 probes were required before the model was found. However, fewer probes are typically required when the data is not based on synthetic models.

Using the connectionist parallel approach, and databases of 512 objects, with 16 points in each object, and scenes of 200 points, we have achieved probe times of under 300 milliseconds on a 32-processor CM-2.

In any case, we are able to perform recognition of models in less than a minute, with dozens of models in the database, which is considerably beyond the state of the art in model-based object recognition, and we are satisfied that real-time recognition using these methods is possible with the appropriate parallel hardware.

4. Motion parameter estimation from optical flow

This work is considerably unrelated to the Statement of Work of the contract, and was only lightly supported by the contract. However, because student V. Sundareswaran became interested in analyzing sensor motion, and because we wished to develop global methods that support a greater degree of stability than existing algorithms, we pursued the topic of motion parameter estimation.

This is a problem that has concerned the IU community for more than a decade. Immense amounts of effort have been expended on motion analysis. There are two aspects to the problem. First, given an image sequence, we typically wish to understand the image motion from pixel to pixel. The result is a vector field, called the image flow field or the optical flow field. The idea is that every pixel points to the corresponding pixel in the next frame, which moves according to the motion of the sensor (and potentially the motion of the object imaged at the pixel). Second, there is the problem of determining the motion parameters from the optical flow field. This is an inverse problem, in that the optical flow is determined by the instantaneous translation and rotation of the sensor (and objects). In addition, there is a dependence on the distance from the object to the sensor. Accordingly, one can hope to reconstruct the motion (translation and rotation) parameters, and the depths to the objects.

What has interested most researchers in this topic is the depth dependence. Theoretically, one can determine the optical flow, analyze the result mathematically, and determine the parameters of motion of the sensor, the depth to scene points, and identify moving objects. The last two objectives are useful for image segmentation and for object identification, both of which are fundamental to ATR applications.

Alas, all of the algorithms are terribly unstable. There is some hope of being able to segment out moving objects, using work recently developed at Sarnoff Research Labs, but the task of extracting depths from image motion flow alone, using any of the existing algorithms, is completely hopeless. The lack of stability is not usually admitted or discussed prominently, but is evident from the methods. The difficulty is that the depths must be determined from local data, whereas the optical flow field can only be determined approximately, sparsely, (for example where there are moving point features), and only with noise. If one is to have any hope of making use of optical flow field data, then the algorithm must be global.

Globally, only the motion parameters can be determined. However, once motion parameters for the sensor are determined, it should be considerably easier to extract moving objects (providing they are small, and thus have not significantly degraded the motion parameter estimation), and the depths. Accordingly, we have developed global methods for motion parameter estimation from optical flow fields.

We developed two methods for the determination of the parameters of motion of a sensor given the vector flow field induced by an imaging system governed by a perspective transformation of a rigid scene. We assume that the flow field $V = (u(x,y),v(x,y))$ is given. Both algorithms are new, and both are extremely simple. The first algorithm is called the *flow circulation algorithm*, and determines the rotational parameters. It uses the curl of the flow field, curl(V). We do not present details here, but we show in our papers that under many conditions, the curl is approximately a linear function of the form $g(x,y) = ax + by + c$ (where x and y are image coordinate values). The coefficients of the linear function, a, b, and c, which may be determined by simple regression, are proportional to the desired rotational parameters of motion. Circulation values may be used in place of curl values, resulting in less noise.

The second algorithm, the *FOE search algorithm*, determines the translational parameters of the motion independently of the first algorithm. This algorithm extends a recently introduced method of Heeger and Jepson, giving a method for searching for the image focus of expansion. Specifically, for every location (x_0,y_0) in the image plane, we compute the function $u\cdot(-y+y_0) + v\cdot(x-x_0)$, which is the component of flow orthogonal to the radial direction scaled by the distance from the point. When (x_0,y_0) is located at the focus of expansion, this function will be a quadratic polynomial (of a special form). The algorithm determines for every (x_0,y_0) whether the computed function is a quadratic polynomial, in order to locate the focus of expansion. We suggest several candidate methods for determining when the function has the appropriate form; one method involves filtering the function by a collection of of circular-surround zero-mean receptive fields. The other methods project the function onto a linear space of quadratic polynomials, and measures the distance between the two functions. Each suggested method can be formulated as the evaluation of a quadratic functional at each candidate point, and the first two methods permit a quadratic search for the focus of expansion, yielding an especially rapid search.

It is this final point, the quadratic nature of the error surface, that is the most startling result of this work. The fact that there exists a global method for the determination of motion parameter is not a surprise — it is only a surprise that these methods had not been previously developed and exploited. However, the simplicity and reduction in cost due to projecting onto a space of quadratic polynomials are great advantages of the method, and argue strongly for its viability as part of a motion analysis system.

As an example, we show a contour plot (together with gradient vectors) for the polynomial projection method for the FOE search algorithm applied to a synthetic flow field (see Figure 3).

Either algorithm may be used in isolation to obtain some parameters, from which the other parameters may then be easily determined. However, the two algorithms in combination provide multiple pathways for independent determination of motion parameters. In the paper accepted for publication in *PAMI* (reference [13]), we demonstrate the viability of both approaches with a number of calculations for typical situations, and a number of experiments using synthetically-generated motion data. Presentation of other details of the methods are given in references [11,12,14]. Sundareswaran has since applied the methods to real motion data sequences, and has investigated sources of error and noise in the system. We are currently in touch with NASA-Ames researchers who will supply us with additional data.

Follow-up work would attempt to make use of the motion parameter estimation methods to locate moving objects, or to estimate range values, so as to determine (for example) ground plane orientation.

5. Uncertainty reasoning

Continuing a longstanding interest in uncertainty reasoning, Professor Hummel has devoted some of the time supported by the AFAL contract to pursue theories of information fusion and uncertainty calculi.

Although only one conference paper was presented (reference [15]), a major journal article is in preparation, and will be submitted soon. All the work is joint with Professor Larry Manevitz, at Haifa University.

Reasoning with uncertainty is a field with many different approaches and viewpoints, with important applications to sensor design and autonomous system development. We attempt to unify some of the different approaches by introducing a common philosophical framework under which different calculi may be developed. Each calculus reflects different design choices compatible with the philosophical tenets. The tenets postulate that uncertainty in representations can be viewed as a degree of dispersion of opinions, and that the space of opinions operates as a separate sample space distinct from the underlying sample space on which probabilities are normally defined. Different calculi result when choices are made for the representation of the opinions, the method for combining opinions, the method for juxtaposing multiple sets of opinions, and the way of measuring the spread in the opinions.

Many sensor fusion systems make use of an evidential reasoning system, where evidence is combined with current measurements in order to maintain states of belief and confidence in a set of hypotheses. These systems are all motivated by the fact that the sensor is supposed to provide more than

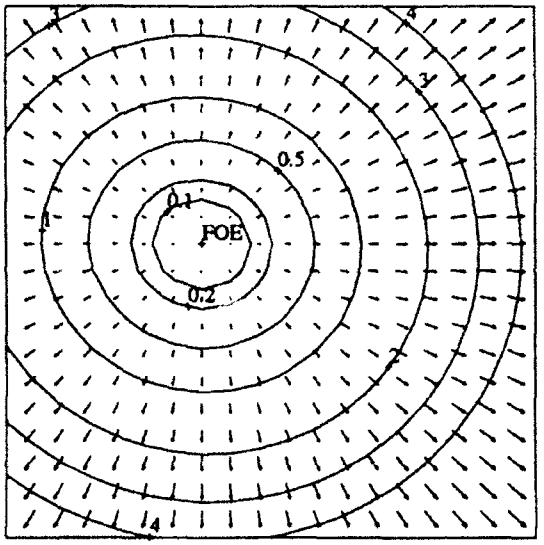


Figure 3. Contour and gradient plot of the error function $E(x_0,y_0)$ obtained using the quadratic polynomial projection method for the image flow field induced by a translational velocity of (5,2,20) and rotational velocity of (0.1,0.025,–0.05) when imaging a corridor (four side walls and a back wall). The focus of expansion is correctly located as the zero of the error surface. It can be seen that the error surface is quadratic, and thus the zero can be found easily.

simple measurements, but also confidence levels in the measurements, and ultimately confidence levels in propositions and hypotheses developed from those measurements. The fundamental concepts always involve quantities related to the degree of validity of a proposition, such as a probability, and other quantities related to the degree of certainty in the assertion of the degree of belief. Various calculi are used for representing these concepts and performing the calculations, including Bayesian networks, fuzzy logic, Kalman filtering, and the Dempster/Shafer theory of evidence. Each calculus has certain theoretical underpinnings, although a universally accepted methodology is still lacking.

Our approach, and the unified treatment of calculi, is founded on certain fixed tenets. The tenets are:

(1) That uncertainty can be represented by a distribution of opinions, whereas certainty is represented by unanimity of multiple opinions.

(2) By an "opinion," we mean an estimate of a quantity that is functionally related to a (frequency-based) probability, or is a subjective estimate of a conditional probability, or is an estimate of a well-defined quantity representing the likelihood of a given proposition based upon given evidence.

(3) When opinions are combined in order to make estimates that are conditioned on combinations of evidence, a precise and well-defined combination formula should be used. In the case that the opinions are estimates of probabilities, Bayes' rule should be used to combine pairs of opinions to

yield a new opinion.

The propositions can relate to measurement values obtained from different sensors, or can be propositions that are developed and based upon those measurements (such as the presence or absence of a target class). For ATR applications, we might be combining information from different modalities, or from different target recognition algorithms, or from different environment variables together with feature data.

A variety of different calculi are obtainable from these philosophical foundations, depending upon the values that the opinions are supposed to represent, and depending upon the assumptions used in the updating process. Each design choice leads to a different calculus, giving a Chinese menu of uncertainty calculi. Different applications will require different design choices. One of the calculi that arises is precisely the Dempster/Shafer calculus. Another is the simplified Kalman filter.

We present here the formulas for yet another calculus that arises in our uniform framework. We envision a collection of "experts" E, with each expert ω in E expressing an opinion $x_\omega(\lambda)$ for every label $\lambda \in \Lambda$. Suppose that the opinions are represented by the logarithms of values determined by the probabilities for a particular proposition. The values, however, are not simply the logarithms of estimated conditional probabilities based on the known observations, but rather the log of the ratio of the conditional probability and the prior probability. Specifically, we set:

$$x_\omega(\lambda) \approx \log\left[\frac{\text{Prob}(\lambda \mid \text{Information})}{\text{Prob}(\lambda)}\right],$$

where the "Information" is the information shared by the experts in E, and the denominator is the prior probability of label λ in the absence of any information. Note that these probabilities are defined over the usual sample space of problem instances, and not the set of experts. The equations are approximate, because each expert is making an estimate. These values are the representation suggested by Charniak for probabilistic reasoning. Each expert will have a different estimate of this log-ratio, and the statistics that we propose to maintain are the mean and variance of these values. Thus if an expert regards a proposition as being 4 times as likely due to the given measurements as opposed to its probability in the absence of information, then the expert contributes log(4) as the opinion in the set of opinions, from which we measure the mean and variance. Note that if the information has no influence on the prior probability, according to an expert, then the expert's opinion will be log(1), i.e., zero.

The state of the system is thus represented by two vectors:

$$\mu(\lambda) = \underset{\omega \in E}{\text{Avg}}(x_\omega(\lambda)),$$

and:

$$\sigma(\lambda) = \left[\underset{\omega \in E}{\text{Avg}}(x_\omega(\lambda) - \mu(\lambda))^2\right]^{1/2}.$$

Unlike the Dempster/Shafer representation, which requires 2^N values for the specification of the state, this formulation requires only $2N$ values, where N is the number of possible labels (i.e., $N = \#\Lambda$).

Now, if two such sets of opinions are to be combined, we chose to take the set product of the opinions. If a composite opinion is to be formed from two individual opinions, and if we may suppose a conditional independence between the information sources on which the two experts are basing their opinions, then it can be shown that modulo a uniform additive constant, the two opinions may be summed. This comes from Bayes formula, using conditional independence of the information sources, yielding:

$$\frac{\text{Prob}(\lambda \mid \text{Info}_1, \text{Info}_2)}{\text{Prob}(\lambda)} = \alpha \frac{\text{Prob}(\lambda \mid \text{Info}_1)}{\text{Prob}(\lambda)} \cdot \frac{\text{Prob}(\lambda \mid \text{Info}_2)}{\text{Prob}(\lambda)}$$

The two sides are equal, except for a proportionality constant, which is independent of λ. Taking the log of both sides, we see that it is logical to set:

$$x_{(\omega_1,\omega_2)}(\lambda) = x_{\omega_1}(\lambda) + x_{\omega_2}(\lambda).$$

The proportionality constant has been dropped, which means that the opinions can be off by a constant additive amount (but the same constant for all the labels λ). This skewing, it turns out, is unimportant, since we are only concerned in the relative size of the components over different labels λ. This then is the updating method, and we can see that independence (conditioned on every label) is required for the information. A mathematical statement of the independence assumptions says that

$$\text{Prob}(\text{Info}_1 \mid \text{Info}_2, \lambda) = \text{Prob}(\text{Info}_1 \mid \lambda)$$

for all λ in Λ. These are strong requirements, but potentially valid in some circumstances.

Using the product formulation for obtaining the the combination set of opinions, we find that the following formulas hold. The mean and standard deviation of the opinions of the experts in set E_1 are denoted by (μ_1,σ_1), and the opinions of the experts in set E_2 give rise to the state (μ_2,σ_2). Then

$$\mu(\lambda) = \mu_1(\lambda) + \mu_2(\lambda),$$
$$\sigma(\lambda) = \left[\sigma_1^2(\lambda) + \sigma_2^2(\lambda)\right]^{1/2}.$$

That is, the resulting calculus can be specified by the statement that mean log-probability opinions should be added, and variances also add. We look for a situation where the resulting mean opinion is either very large positive, or very large negative, with a relatively tight variance (much less than the magnitude of the opinion) in order to conclude that the corresponding proposition is true or false.

6. Wavelet slice theorem and variants

Our work on wavelet methods for SAR image formation has, to date, been mostly based on mathematical explorations. The work was begun by postdoctoral student Jacques Froment, and has been taken up by mathematics graduate student Jordan Mann. We announce here principal results. The first two results are termed "Wavelet Slice Theorems," and they form wavelet analogs of the "Projection Slice Theorem," which forms the basis of the SAR technology (as well as tomography technology). The last two results are the most recent work, and suggest exciting new methods for practical SAR image formation that might provide practical advantages over current methods.

The problem of reconstructing a function of n variables from its integrals along hyperplanes was first studied by Radon. The problem, its two-dimensional version in particular, has found applications in such diverse fields as radio astronomy, molecular biology, and computer-assisted tomography. More recently the foundational mathematics has been applied to synthetic aperture radar (SAR).

We have proposed several new formulae which hold particular promise for SAR, since they take advantage of the convolution inherent in the transmission and reception of the radar signal. The first two formulae are based on arbitrary wavelets and thus open the possibility of determining, analytically or empirically, the wavelet or wavelets most effective for a given application. Notably, finely tuned orthogonal wavelets are not needed in our methods.

Since these ideas are as yet unpublished, but may lead to exciting new methods for SAR image formation, we present a mathematical summary of our work below.

We consider a function $f(x,y)$, and define $P_\theta f(t)$ to be the Radon transform of the function at angle θ, given by

$$P_\theta f(t) = \int_{-\infty}^{\infty} f(t\cos\theta - \tau\sin\theta, t\sin\theta + \tau\cos\theta)\, d\tau.$$

For any particular t, $P_\theta f(t)$ is the integral of the function along a specified line in the coordinate plane.

Given a function $f(x)$ of one variable and a 1-D wavelet $\psi(x)$, we define $W_s f(u)$, the wavelet transform of f at location u and scale s by:

$$W_s f(u) = (f * \tilde{\psi}_s)(u),$$

the convolution of the signal f and the wavelet, and where:

$$\psi_s(x) = \sqrt{s}\,\psi(sx), \text{ and}$$

$$\tilde{\psi}_s(x) = \psi_s(-x).$$

Given a function $f(x,y)$ of two variables and a 2-D wavelet $\psi(x,y)$, the wavelet transform $W_s f(u,v)$ of f at location (u,v) and at the scale s, is defined by:

$$W_s f(u,v) = (f * \tilde{\psi}_s)(u,v),$$

where $\psi_s(x,y)$ and $\tilde{\psi}$ are defined similarly as above, but for two variables.

In spotlight mode SAR, if we "chirp" a function $h(t)$ in a direction θ toward an image $g(x,y)$, the return signal, in a certain sense, is essentially $h * P_\theta g(t)$. Accordingly, in the following results, we consider $g(x,y)$ to be the function for reconstruction.

Result 1. We have the following formula:

$$C\int_{-\pi}^{\pi}\int_0^{+\infty} W_s P_\theta g(U(\theta))\frac{1}{\sqrt{s}}\,ds\,d\theta = (m_{(1/r)} * \tilde{g})(x,y)$$

where C is a multiplicative constant, $m_{(1/r)}(x,y)$ is the function whose Fourier transform is $1/\sqrt{\xi^2+\eta^2}$, and $U(\theta) = x\cos\theta + y\sin\theta$.

Since the convolution on the right side of the equation can be inverted, we can recover g. In fact, we have:

$$g(x,y) = C\int_{-\pi}^{\pi}\int_0^{+\infty} H' W_s P_\theta g(U(\theta))\frac{1}{\sqrt{s}}\,ds\,d\theta,$$

where H' denotes differentiation followed by the Hilbert transform.

The significance of this result is that we may use any 1-D wavelet to construct $W_s P_\theta g$ by chirping a collection of different signals at varying scales at each angle. The function g may then be recovered through the indicated formula, which involves a Hilbert transform.

Compared with typical SAR analysis, a rasterization onto a coordinate grid is not necessary, since all the processing is 1-D before integrating over space. However, like regular SAR processing, FFTs (or this case, Hilbert transforms) are necessary.

Result 2. Let ψ be any 2-D wavelet. By the wavelet transform inversion formula:

$$g(x,y) = \frac{1}{C_\psi}\int_0^{\infty} s\cdot(\psi_s * W_s g(x,y)\,ds \tag{1}$$

where C_ψ is a constant associated with ψ. Writing the convolution on the right side as the inverse Fourier Transform of its own Fourier transform, and the convolution theorem, we find the the right side of (1) is equal to:

$$\frac{1}{2\pi}\frac{1}{C_\psi}\int_0^{\infty} s\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\psi}_s(\zeta,\xi)\hat{\tilde{\psi}}_s(\zeta,\xi)\hat{g}(\zeta,\xi)e^{i(x\zeta+y\xi)}\,d\zeta\,d\xi\,ds.$$

The integral over $d\zeta d\xi$ may be easily transformed into a polar coordinate integral $/\rho/d\rho d\theta$.

By making the change of variables, and regrouping terms, (1) may be rewritten as follows. Let:

$$C_{s,\theta}(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \hat{\psi}_s(\rho\cos\theta,\rho\sin\theta)\,\hat{\tilde{\psi}}_s(\rho\cos\theta,\rho\sin\theta)\,/\rho/\,e^{i\rho u}\,d\rho.$$

Then (1) becomes, after some algebra:

$$g(x,y) = \frac{1}{2\pi}\frac{1}{C_\psi}\int_0^{\infty} s\cdot\int_0^{\pi}(P_\theta g * C_{s,\theta})(x\cos\theta + y\sin\theta)\,d\theta\,ds \tag{2}$$

We see that g can be reconstructed using an extremely simple formula, providing we have available the return signals from a spotlight mode radar that has emitted a collection of $C_{s,\theta}$ signals. This leads us to investigate properties of $C_{s,\theta}$.

Suppose we set $C_\theta(u)$ to be equal to $C_{s,\theta}(u)$ when $s = 1$. Then is is not hard to show that:

$$C_{s,\theta}(u) = C_\theta(su), \text{ and}$$

$$C_\theta(-u) = C_\theta(u).$$

In the case when ψ is real (and real-valued wavelets are the normal situation), we have:

$$C_\theta(u) = \sqrt{\frac{2}{\pi}} \int_0^\infty |\hat{\psi}(\rho\cos\theta, \rho\sin\theta)| \rho\cos(\rho u)\, d\rho.$$

Basically, $C_{s,\theta}(u)$ is a function whose Fourier Transform is $\rho \cdot |\hat{\psi}_s|^2(\rho\cos\theta, \rho\sin\theta)$ as a function of ρ. Accordingly, we expect that for a suitable choice of ψ, the function $C_{s,\theta}$ will be smooth and essentially compactly supported, and thus a good choice for a chirp signal.

The problem with Result 2 is that an entire collection of different scales are needed at each angle. The same situation occurs, of course, with standard SAR. The issue is whether there exists a method to chirp a range of different frequencies, and sort out the necessary coefficients, as is done with standard SAR processing. Amazingly, it is as simple, if not simpler, in the wavelet case.

Result 3. Let us suppose that g is wavelet-domain band-limited, by which we mean that there exists $T>0$ such that for all $s>T$, $W_s g(u,v) = 0$ for all (u,v). This is conceptually similar to the familiar assumption of band-limitedness. Then in formula (2) above, the integration with respect to s can be truncated at T instead of going to infinity, and it is possible to change the order of integration and integrate with respect to s first. Integrating with respect to s first, but making the change of variable $u=sx$, we find that:

$$g(x,y) = \frac{1}{2\pi} \frac{1}{C_\psi} \int_0^\pi (P_\theta g * h_\theta)(U(\theta))\, d\theta, \tag{3}$$

where:

$$h_\theta(v) = \frac{1}{v^2} \int_0^{vT} u \cdot C_\theta(u)\, du.$$

That is, rather than needing a collection of returns, it suffices to filter $P_\theta g$ by h_θ, which is the return at angle θ that is obtained by chirping h_θ.

It can be shown that $h_\theta(v)$ is continuous at 0, and with proper choice of wavelet it can be made to decay like $1/v^2$. Thus it is entirely reasonable to expect that h_θ can be implemented as a valid chirp.

We see that once the convolution has been performed, the only other computation needed is integration with respect to θ. Note that the desired chirp $h_\theta(v)$ depends on θ, so that potentially different signals are needed for each position along the flight path. However, if the wavelet ψ is radially symmetric, then C_θ and h_θ will be independent of θ. Although radially symmetric wavelets are known to exist, it is not known if these wavelets have the other desirable properties required to make h_θ a desirable chirp function.

Our final result is a generalization, and returns us to the domain of Fourier analysis, as opposed to wavelet analysis. The result puts the previous result in perspective, showing the relationship between existing reconstruction methods, and the proposed new method.

Result 4. The standard reconstruction formula for tomography, and implicitly for SAR image formation, is:

$$g(x,y) = \frac{1}{2\pi} \int_0^\pi (H' P_\theta g)(x\cos\theta + y\sin\theta)\, d\theta.$$

The wavelet form of the same equation was given in Result 1. Here, the operator H' denotes differentiation followed by the Hilbert transform. For arbitrary functions, H' cannot be implemented as a convolution filter in the spatial domain, and must be achieved by multiplication by $|\rho|$ in the Fourier domain. However, if we suppose that f is band-limited in the usual sense, then application of H' to f can be achieved by convolution with any function q, where $q(x)$ is a function whose Fourier transform equals $|\rho|$ in support in the spectral domain of f. That is:

$$g(x,y) = \frac{1}{2\pi}\int_0^{\pi} (P_\theta g * q)(U(\theta))\, d\theta, \tag{4}$$

where U is defined as before, and it is assumed that $P_\theta g$ is bandlimited (as a function of u), and q is suitably defined. This result is known (for example, see A. Kak, "Computerized tomography with x-ray, emission, an ultrasound sources," *Proc. IEEE*, September 1979), but not widely utilized. We see, however, that Eqn. (4) has the same form as Eqn. (3), and implies a simple method for SAR image formation, as we discuss below. However, the assumptions are different: Eqn. (3) is based on wavelet-bandlimitedness of g, which is a form of local-frequency limitation, whereas Eqn. (4) assumes the usual bandlimitedness of $P_\theta g$ for all θ.

We now discuss the implementation of formulas of the form (3) and (4) for SAR image formation. These formulas are very well-suited to SAR for the following reason. All existing Radon transform reconstruction techniques assume that the values of the Radon transform $P_\theta g(x)$ or its Fourier transform $\hat{P}_\theta g(\rho)$ are known for at least some values of θ and x or ρ. In SAR, however, this is not really the case; the radar signal received from the reflecting object is not $P_\theta g$ but $P_\theta g * S$, where S is is the transmitted signal. Using a filter, it is possible to reconstruct $\hat{P}_\theta g(\rho)$ for some values of θ and ρ from these values, which can then be rerasterized to a coordinate grid, in order to obtain $\hat{g}(\xi,\eta)$, and then $g(x,y)$. Clearly, considerable processing is required on the return signals, and the data $P_\theta g * S$ are data that are used for the reconstruction, and not directly related to the image data in an untransformed state. With our methods, however, convolution of $P_\theta g$ with another function, which we may call the convolving function, is the first step in the reconstruction technique. We therefore recommend implementation of any of these three methods by using its convolving function as the transmitted signal. In this way, the convolution of $P_\theta g$

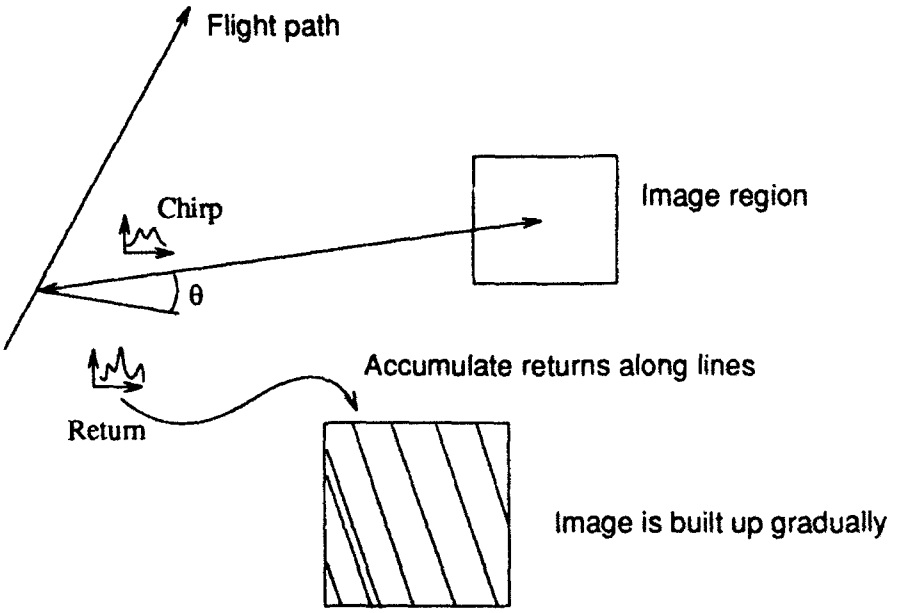


Figure 4. An outline of the computations required for each point of data collection in the SAR image formation method suggested by Eqns. (3) and (4). For each value of θ, the return signal forms a line of data, which is spread through an array of accumulators in a direction perpendicular to θ.

first step in the reconstruction technique. We therefore recommend implementation of any of these three methods by using its convolving function as the transmitted signal. In this way, the convolution of $P_\theta g$ with the transmitted signal is no longer a nuisance that must be reversed, but the first step in the solution!

Accordingly, the method for reconstruction is as follows. At each point along the flight path (see Figure 4), a convolution function $h_\theta(t)$ is sent, and the return signal $f_\theta(u)$ is recorded. For any given (x,y), the contribution of f_θ toward the reconstructed value of $g(x,y)$ depends on $f_\theta(x\cos\theta+y\sin\theta)$. That is, we have the formula:

$$g(x,y) = \int_0^{2\pi} f_\theta(x\cos\theta+y\sin\theta)\, d\theta.$$

Note that the contribution for any given $f_\theta(u)$ is constant along a line perpendicular to the direction $(\cos\theta,\sin\theta)$. Thus, for each appropriate value of u, an accumulator along a line perpendicular to the direction θ is incremented by an amount $f_\theta(u)$. This can be done in parallel for different values of u.

To summarize, the suggested algorithm is as follows:

Initialize $g(x,y)=0$ for all (x,y).

For all θ in the range of available angles, do:

- Obtain a return signal $f_\theta(u)$ using a chirp $h_\theta(t)$.
- For each value of u, do

  Use Bresenham's algorithm to walk along a line in a direction perpendicular to θ, (the line $U(\theta) = u$, and increment an each accumulator by a portion of $f_\theta(u)$.

  The portion should be computed using standard anti-aliasing methods.

We see that standard computer graphics techniques can be used to form the computations. The algorithms are easily parallelized. The total complexity for each point of θ is proportional to the number of pixels in the image. With appropriate hardware, the accumulation could take place during the collection of data, so that the final scene would be available by the time the final data point is collected.

U.S. GOVERNMENT PRINTING OFFICE 750-113